Hermitian operators and convex functions

Jean-Christophe Bourin

E-mail: bourinjc@club-internet.fr

Université de Cergy-Pontoise, dépt. de Mathématiques

2 rue Adolphe Chauvin, 95302 Pontoise, France

Abstract.

We establish several convexity results for Hermitian matrices. For instance: Let A, B be Hermitian and let f be a convex function. If X and Y stands for $f(\{A + B\}/2)$ and $\{f(A) + f(B)\}/2$ respectively, then there exist unitaries U, V such that

$$X \leq \frac{UYU^* + VYV^*}{2}.$$

This is nothing but the matrix version of the scalar convexity inequality

$$f\left(\frac{a + b}{2}\right) \leq \frac{f(a) + f(b)}{2}.$$

As a consequence we get, $\lambda_{2j-1}(X) \leq \lambda_j(Y)$, where $\lambda_j(\cdot)$ are the eigenvalues arranged in decreasing order.

Keywords: Hermitian operators, eigenvalues, operator inequalities, Jensen's inequality
Mathematical subjects classification: 47A30 47A63

Introduction

The main aim of this paper is to give a matrix version of the scalar inequality

$$f\left(\frac{a + b}{2}\right) \leq \frac{f(a) + f(b)}{2} \tag{1}$$

for convex functions f on the real line.

Capital letters A, $B \ldots Z$ mean n-by-n complex matrices, or operators on a finite dimensional Hilbert space $\mathcal{H}$; I stands for the identity. When A is positive semidefinite, resp. positive definite, we write $A \geq 0$, resp. $A > 0$.

A classical matrix version of (1) is von Neuman's Trace Inequality: For Hermitians A, B,

$$\operatorname{Tr} f\left(\frac{A + B}{2}\right) \leq \operatorname{Tr} \frac{f(A) + f(B)}{2} \tag{2}$$

When f is convex and monotone, we showed [2] that (2) can be extended to an operator inequality: There exists a unitary U such that

$$f\left(\frac{A + B}{2}\right) \leq U \cdot \frac{f(A) + f(B)}{2} \cdot U^* \tag{3}$$

1

We also established similar inequalities involving more general convex combinations. These inequalities are equivalent to an inequality for compressions. Recall that given an operator Z and a subspace $\mathcal{E}$ with corresponding orthoprojection E, the compression of Z onto $\mathcal{E}$, denoted by $Z_\mathcal{E}$, is the restriction of EZ to $\mathcal{E}$. Inequality (3) can be derived from: For every Hermitian A, subspace $\mathcal{E}$ and monotone convex function f, there exists a unitary operator U on $\mathcal{E}$ such that

$$f(A_\mathcal{E}) \leq U f(A)_\mathcal{E} U^*. \tag{4}$$

Inequalities (3) and (4) are equivalent to inequalities for eigenvalues. For instance (4) can be rephrased as

$$\lambda_j(f(A_\mathcal{E})) \leq \lambda_j(f(A)_\mathcal{E}), \quad j = 1, 2, \ldots$$

where $\lambda_j(\cdot)$, $j = 1, 2, \ldots$ are the eigenvalues arranged in decreasing order and counted with their multiplicities. Having proved an inequality such as (3) for monotone convex functions, it remains to search counterparts for general convex functions. We derived from (3) the following result for even convex functions f: Given Hermitians A, B, there exist unitaries U, V such that

$$f\left(\frac{A+B}{2}\right) \leq \frac{U f(A) U^* + V f(B) V^*}{2}. \tag{5}$$

This generalizes a wellknown inequality for the absolute value,

$$|A + B| \leq U|A|U^* + V|B|V^*$$

We do not know whether (5) is valid for all convex functions.

In Section 1 we present a counterpart of (4) for all convex functions. This will enable us to give, in Section 2, a quite natural counterpart of (3) for all convex functions. Though (3) can be proven independently of (4) -and the same for the counterparts-, we have the feeling that in the case of general convex functions, the approach via compressions is more illuminating.

1. Compressions

Our substitute to (4) for general convex functions (on the real line) is:

Theorem 1.1. *Let A be Hermitian, let $\mathcal{E}$ be a subspace and let f be a convex function. Then, there exist unitaries U, V on $\mathcal{E}$ such that*

$$f(A_\mathcal{E}) \leq \frac{U f(A)_\mathcal{E} U^* + V f(A)_\mathcal{E} V^*}{2}.$$

Consequently, for $j = 1, 2, \ldots,$

$$\lambda_{2j-1}(f(A_\mathcal{E})) \leq \lambda_j(f(A)_\mathcal{E}).$$

Proof. We may find spectral subspaces $\mathcal{E}'$ and $\mathcal{E}''$ for $A_\mathcal{E}$ and a real r such that

   (i) $\mathcal{E} = \mathcal{E}' \oplus \mathcal{E}''$,

   (ii) the spectrum of $A_{\mathcal{E}'}$ lies on $(-\infty, r]$ and the spectrum of $A_{\mathcal{E}''}$ lies on $[r, \infty)$,

   (iii) f is monotone both on $(-\infty, r]$ and $[r, \infty)$.

Let k be an integer, $1 \leq k \leq \dim \mathcal{E}'$. There exists a spectral subspace $\mathcal{F} \subset \mathcal{E}'$ for $A_{\mathcal{E}'}$ (hence for $f(A_{\mathcal{E}'})$), $\dim \mathcal{F} = k$, such that

$$\begin{aligned}
\lambda_k[f(A_{\mathcal{E}'})] &= \min_{h \in \mathcal{F};\ \|h\|=1} \langle h, f(A_\mathcal{F})h \rangle \\
&= \min\{f(\lambda_1(A_\mathcal{F}))\,;\ f(\lambda_k(A_\mathcal{F}))\} \\
&= \min_{h \in \mathcal{F};\ \|h\|=1} f(\langle h, A_\mathcal{F} h \rangle) \\
&= \min_{h \in \mathcal{F};\ \|h\|=1} f(\langle h, A h \rangle)
\end{aligned}$$

where at the second and third steps we use the monotony of f on $(-\infty, r]$ and the fact that $A_\mathcal{F}$'s spectrum lies on $(-\infty, r]$. The convexity of f implies

$$f(\langle h, Ah \rangle) \leq \langle h, f(A)h \rangle$$

for all normalized vectors h. Therefore, by the minmax principle,

$$\begin{aligned}
\lambda_k[f(A_{\mathcal{E}'})] &\leq \min_{h \in \mathcal{F};\ \|h\|=1} \langle h, f(A)h \rangle \\
&\leq \lambda_k[f(A)_{\mathcal{E}'}].
\end{aligned}$$

This statement is equivalent (by unitay congruence to diagonal matrices) to the existence of a unitary operator U_0 on $\mathcal{E}'$ such that

$$f(A_{\mathcal{E}'}) \leq U_0 f(A)_{\mathcal{E}'} U_0^*.$$

Similarly we get a unitary V_0 on $\mathcal{E}''$ such that

$$f(A_{\mathcal{E}''}) \leq V_0 f(A)_{\mathcal{E}''} V_0^*.$$

Thus we have

$$f(A_\mathcal{E}) \leq \begin{pmatrix} U_0 & 0 \\ 0 & V_0 \end{pmatrix} \begin{pmatrix} f(A)_{\mathcal{E}'} & 0 \\ 0 & f(A)_{\mathcal{E}''} \end{pmatrix} \begin{pmatrix} U_0^* & 0 \\ 0 & V_0^* \end{pmatrix}.$$

Besides we note that, still in respect with the decomposition $\mathcal{E} = \mathcal{E}' \oplus \mathcal{E}''$,

$$\begin{pmatrix} f(A)_{\mathcal{E}'} & 0 \\ 0 & f(A)_{\mathcal{E}''} \end{pmatrix} = \frac{1}{2}\left\{ \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} f(A)_\mathcal{E} \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} + \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} f(A)_\mathcal{E} \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \right\}.$$

So, letting

$$U = \begin{pmatrix} U_0 & 0 \\ 0 & V_0 \end{pmatrix} \quad \text{and} \quad V = \begin{pmatrix} U_0 & 0 \\ 0 & -V_0 \end{pmatrix}$$

we get

$$f(A_\mathcal{E}) \leq \frac{U f(A)_\mathcal{E} U^* + V f(A)_\mathcal{E} V^*}{2}. \tag{6}$$

It remains to check that (6) entails

$$\lambda_{2j-1}(f(A_{\mathcal{E}})) \leq \lambda_j(f(A)_{\mathcal{E}}).$$

This follows from the forthcoming elementary observation.   □

Proposition 1.2. *Let X, Y be Hermitians such that*

$$X \leq \frac{UYU^* + VYV^*}{2} \tag{7}$$

for some unitaries U, V. Then, for $j = 1, 2, \ldots$,

$$\lambda_{2j-1}(X) \leq \lambda_j(Y).$$

Proof. By adding a rI term, for a suitable scalar r, both to X and Y, it suffices to show that

$$\lambda_{2j-1}(X) > 0 \quad \implies \quad \lambda_j(Y) > 0. \tag{8}$$

We need the following obvious fact: Given Hermitians A, B,

$$\mathrm{rank}(A + B)_+ \leq \mathrm{rank}A_+ + \mathrm{rank}B_+$$

where the subscript $+$ stands for positive parts. Applying this to $A = UYU^*$ and $B = VYV^*$ we infer that the negation of (8), that is $\lambda_{2j-1}(A + B) > 0$ and $\lambda_j(A)\,(= \lambda_j(B)) \leq 0$, can not hold. Indeed, the relation

$$\mathrm{rank}(A + B)_+ \geq 2j - 1 > (j - 1) + (j - 1) \geq \mathrm{rank}A_+ + \mathrm{rank}B_+$$

would contradict the previous rank inequality.   □

Remark 1.3. From inequality (7) one also derives, as a straightforward consequence of Fan's Maximum Principle [1, Chapter 4],

$$\sum_{j=1}^{k} \lambda_j(X) \leq \sum_{j=1}^{k} \lambda_j(Y)$$

for $k = 1, 2, \ldots$.
  Inequality (7) also implies

$$\lambda_{i+j+1}(X) \leq \frac{1}{2}\{\lambda_{i+1}(Y) + \lambda_{j+1}(Y)\}$$

for i, $j = 0, 1, \ldots$. It is a special case of Weyl's inequalities [1, Chapter 3].

Remark 1.4. For operators acting on an infinite dimensional (separable) space, the main inequality of Theorem 1.1 is still valid at the cost of an additional rI term in the RHS, with $r > 0$ arbitrarily small. See [3, Chapter 1] for the analogous result for (4).

Obviously, for a concave function f, the main inequality of Theorem 1.1 is reversed. But the following is open:

Question 1.5. Let g be a concave function, let A be Hermitian and let $\mathcal{E}$ be a subspace. Can we find unitaries U, V on $\mathcal{E}$ such that

$$g(A)_{\mathcal{E}} \leq \frac{Ug(A_{\mathcal{E}})U^* + Vg(A_{\mathcal{E}})V^*}{2} \quad ?$$

2. Convex combinations

The next two theorems can be regarded as matrix versions of Jensen's inequality. The first one is also a matrix version of the elementary scalar inequality

$$f(za) \leq zf(a)$$

for convex functions f with $f(0) \leq 0$ and scalars a and z with $0 < z < 1$.

Theorem 2.1. *Let f be a convex function, let A be Hermitian, let Z be a contraction and set $X = f(Z^*AZ)$ and $Y = Z^*f(A)Z$. Then, there exist unitaries U, V such that*

$$X \leq \frac{UYU^* + VYV^*}{2}.$$

A family $\{Z_i\}_{i=1}^{m}$ is an isometric column if $\sum_{i=1}^{m} Z_i^* Z_i = I$

Theorem 2.2. *Let f be a convex function, let $\{A_i\}_{i=1}^{m}$ be Hermitians, let $\{Z_i\}_{i=1}^{m}$ be an isometric column and set $X = f(\sum Z_i^* A_i Z_i)$ and $Y = \sum Z_i^* f(A_i)Z_i$. Then, there exist unitaries U, V such that*

$$X \leq \frac{UYU^* + VYV^*}{2}.$$

Corollary 2.3. *Let f be a convex function, let A, B be Hermitians and set $X = f(\{A + B\}/2)$ and $Y = \{f(A) + f(B)\}/2$. Then, there exist unitaries U, V such that*

$$X \leq \frac{UYU^* + VYV^*}{2}.$$

Recall that the above inequality entails that for $j = 1, 2, \ldots,$

$$\lambda_{2j-1}(X) \leq \lambda_j(Y).$$

We turn to the proof of Theorems 2.1 and 2.2.

Proof. Theorem 2.1 and Theorem 1.1 are equivalent. Indeed, to prove Theorem 1.1, we may assume that $f(0) = 0$. Then, Theorem 1.1 follows from Theorem 2.2 by taking Z as the projection onto $\mathcal{E}$.

Theorem 1.1 entails Theorem 2.2: to see that, we introduce the partial isometry J and the operator $\tilde{A}$ on $\mathcal{H} \oplus \mathcal{H}$ defined by

$$J = \begin{pmatrix} Z & 0 \\ (I - |Z|^2)^{1/2} & 0 \end{pmatrix}, \quad \tilde{A} = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}.$$

Denoting by $\mathcal{H}$ the first summand of the direct sum $\mathcal{H} \oplus \mathcal{H}$, we observe that

$$f(Z^* A Z) = f(J^* \tilde{A} J) : \mathcal{H} = J^* f(\tilde{A}_{J(\mathcal{H})}) J : \mathcal{H}.$$

where $X : \mathcal{H}$ means the restriction of an operator X to the first summand of $\mathcal{H} \oplus \mathcal{H}$. Applying Theorem 1.1 with $\mathcal{E} = J(\mathcal{H})$, we get unitaries U_0, V_0 on $J(\mathcal{H})$ such that

$$f(Z^* A Z) \leq J^* \frac{U_0 f(\tilde{A})_{J(\mathcal{H})} U_0^* + V_0 f(\tilde{A})_{J(\mathcal{H})} V_0^*}{2} J : \mathcal{H}.$$

Equivalently, there exist unitaries U, V on $\mathcal{H}$ such that

$$\begin{aligned}
f(Z^* A Z) \leq & \frac{U J^* f(\tilde{A})_{J(\mathcal{H})} (J : \mathcal{H}) U^* + V J^* f(\tilde{A})_{J(\mathcal{H})} (J : \mathcal{H}) V^*}{2} \\
= & \frac{1}{2} \left\{ U J^* \begin{pmatrix} f(A) & 0 \\ 0 & f(0) \end{pmatrix} (J : \mathcal{H}) U^* + V J^* \begin{pmatrix} f(A) & 0 \\ 0 & f(0) \end{pmatrix} (J : \mathcal{H}) V^* \right\} \\
= & \frac{1}{2} U \{ Z^* f(A) Z + (I - |Z|^2)^{1/2} f(0) (I - |Z|^2)^{1/2} \} U^* \\
& + \frac{1}{2} V \{ Z^* f(A) Z + (I - |Z|^2)^{1/2} f(0) (I - |Z|^2)^{1/2} \} V^*.
\end{aligned}$$

Using $f(0) \leq 0$ we obtain the first claim of Theorem 2.2.

Similarly, Theorem 1.1 implies Theorem 2.3 (we may assume $f(0) = 0$) by considering the partial isometry and the operator on $\oplus^m \mathcal{H}$,

$$\begin{pmatrix} Z_1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ Z_m & 0 & \cdots & 0 \end{pmatrix}, \qquad \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_m \end{pmatrix}.$$

$\square$

We note that our theorems contain two wellknown trace inequalities [4], [5]:

2.4. Brown-Kosaki: Let f be convex with $f(0) \leq 0$ and let A be Hermitian. Then, for all contractions Z,

$$\operatorname{Tr} f(Z^* A Z) \leq \operatorname{Tr} Z^* f(A) Z.$$

2.5. Hansen-Pedersen: Let f be convex and let $\{A_i\}_{i=1}^m$ be Hermitians. Then, for all isometric column $\{Z_i\}_{i=1}^m$,

$$\operatorname{Tr} f(\sum_i Z_i^* A_i Z_i) \leq \operatorname{Tr} \sum_i Z_i^* f(A_i) Z_i.$$

References

[1] R. Bhatia, Matrix Analysis, Springer, Germany, 1996.

[2] J.-C. Bourin, *Convexity or concavity inequalities for Hermitian operators*, Math. Ineq. Appl. 7 $n^0 4$ (2004) 607-620.

[3] J.-C. Bourin, Compressions, Dilations and Matrix Inequalities, RGMIA monograph, Victoria university, Melbourne 2004 (http://rgmia.vu.edu.au/monograph)

[4] L. G. Brown and H. Kosaki, *Jensen's inequality in semi-finite von Neuman algebras*, J. Operator Theory 23 (1990) 3-19.

[5] F. Hansen G. K. Pedersen, *Jensen's operator inequality*, Bull. London Math. Soc. 35 (2003) 553-564.